

Constructing $2m$ -variable Boolean functions with optimal algebraic immunity based on polar decomposition of $\mathbb{F}_{2^{2m}}^*$

Jia Zheng[†], Baofeng Wu[‡], Yufu Chen[§], Zhuojun Liu[¶]

Abstract

Constructing $2m$ -variable Boolean functions with optimal algebraic immunity based on decomposition of additive group of the finite field $\mathbb{F}_{2^{2m}}$ seems to be a promising approach since Tu and Deng's work. In this paper, we consider the same problem in a new way. Based on polar decomposition of the multiplicative group of $\mathbb{F}_{2^{2m}}$, we propose a new construction of Boolean functions with optimal algebraic immunity. By a slight modification of it, we obtain a class of balanced Boolean functions achieving optimal algebraic immunity, which also have optimal algebraic degree and high nonlinearity. Computer investigations imply that this class of functions also behave well against fast algebraic attacks.

Keywords Boolean functions; Algebraic immunity; Polar decomposition; Balanced; Nonlinearity.

*Partially supported by National Basic Research Program of China (2011CB302400).

[†]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. Email: zhengjia11b@mailucas.ac.cn

[‡]Key Laboratory of Mathematics Mechanization, AMSS, Chinese Academy of Sciences, Beijing 100190, China. Email: wubaofeng@amss.ac.cn

[§]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. Email: yfchen@ucas.ac.cn

[¶]Key Laboratory of Mathematics Mechanization, AMSS, Chinese Academy of Sciences, Beijing 100190, China. Email: zliu@mmrc.iss.ac.cn

1 Introduction

Boolean functions play an important role in symmetric cryptography, especially in the stream ciphers based on linear feedback shift registers (LFSRs). They can be used as building blocks in such key stream generators as filter generator and combiner generator. Due to the existence of different kinds of known attacks to stream ciphers, Boolean functions that are useable should satisfy some main criteria such as balancedness, high algebraic degree, high nonlinearity and optimal algebraic immunity.

The notion of algebraic immunity was introduced in [18] by Meier et al. after the great success of algebraic attacks to such well-known stream ciphers as Toyocrypt and LILI-128 [7]. In fact, the algebraic immunity of a Boolean function f is the smallest possible degree of the nonzero Boolean functions that can annihilate f or $f+1$. If it is not big enough, the multivariate polynomial systems derived from the stream ciphers involving f would be efficiently solved, and hence the secret key can be recovered. This is just the clever idea of the standard algebraic attacks introduced (improved, more definitely) by Courtois and Meier [7]. It can be proved that the best possible value of the algebraic immunity of n -variable Boolean functions is $\lceil \frac{n}{2} \rceil$ [7], thus functions attaining this upper bound are often known as algebraic immunity optimal functions, or OAI functions for short.

After OAI Boolean functions were introduced, the natural question of constructing them was considered in a series of work (see e.g. [4, 8, 12, 13]). But the initial constructions only focused on the criterion of optimal algebraic immunity and did not satisfy other criteria of Boolean functions, so they were just of more interest in theory. Besides, though having optimal algebraic immunity, these functions did not resist fast algebraic attacks (FAAs) well. The technique of fast algebraic attack is improved from the standard algebraic attack, the key point of which is to find low degree multiples of Boolean functions used in the ciphers to be attacked such that their products are of reasonable degree [6]. No progress in constructing Boolean functions having all “good” properties was made until 2008. In their pioneering work, Carlet and Feng proposed an infinite class of balanced Boolean functions which had optimal algebraic immunity, optimal algebraic degree and high nonlinearity [5]. Computer experiments implied that the constructed functions also behaved well against fast algebraic attacks (in fact, very recently this was validated by Liu et al. in theory [14]).

In fact, Carlet and Feng seem to have suggested a principle of constructing

Boolean functions achieving optimal algebraic immunity from finite fields, that is consecutive powers of primitive elements of certain cyclic groups should be involved in the functions' supports, which can promise the utility of BCH bound from coding theory in proving the optimal algebraic immunity of the constructed functions. Following this principle, Tu and Deng tried a new idea and (almost) succeeded. They constructed a class of $2m$ -variable Boolean functions based on the additive decomposition

$$\mathbb{F}_{2^{2m}} = \mathbb{F}_{2^m} \times \mathbb{F}_{2^m} \quad (1)$$

which optimized most of the criteria [21], but had two drawbacks that the optimal algebraic immunity of them could only be proved assuming the correctness of a combinatorial conjecture, and the ability of them resisting fast algebraic attacks is bad [3]. Afterwards, Tang et al. adopted a similar technique, constructing a class of OAI functions which also had other good properties and good immunity against fast algebraic attacks [20] (in fact, this was stated by Tang et al. based on computer experiments firstly and proved by Liu et al. in theory lately [15]). Very recently, Jin et al. found a general construction that could involve Tu and Deng's construction and Tang et al.'s construction as special cases [9]. The optimal algebraic immunity of these functions was proved based on a general conjecture proposed in [20]. In all these constructions of even variable OAI functions, the "certain cyclic group" was chosen to be $\mathbb{F}_{2^m}^*$, the multiplicative group of the finite field \mathbb{F}_{2^m} .

In addition to the additive decomposition (1) of $\mathbb{F}_{2^{2m}}$, we also have a multiplicative decomposition of $\mathbb{F}_{2^{2m}}^*$ like

$$\mathbb{F}_{2^{2m}}^* = \mathbb{F}_{2^m}^* \times U, \quad (2)$$

where U is a cyclic subgroup of $\mathbb{F}_{2^{2m}}^*$ of order $(2^m + 1)$. In fact, instead of multiplicative decomposition, this decomposition is often known as the polar decomposition of $\mathbb{F}_{2^{2m}}^*$, which can be used to construct bent and hyper-bent functions [19], and vectorial Boolean functions achieving high algebraic immunity [16]. By choosing the "certain cyclic group" in Carlet and Feng's principle to be $\mathbb{F}_{2^m}^*$, we propose a new construction of $2m$ -variable OAI Boolean functions based on the polar decomposition (2) in this paper, which can be viewed as a multiplicative analog of Tu and Deng's construction. After modifying these functions to be balanced ones, we obtain Boolean functions satisfying almost all main criteria and potentially behaving well against fast algebraic attacks (by potentially we mean that this is only supported by

computational evidence up to present). A big difference in the “modifying to be balanced” process between our construction and the former ones is that, something should be subtracted from the supports of the initially constructed functions since they are “fatter” than that of balanced functions in our construction, while something should be added to the supports of the initially constructed functions since they are “thinner” than that of balanced functions in the former constructions.

The rest of the paper is organized as follows. In Section 2, we give the necessary preliminaries concerning Boolean functions. In Section 3, we prove a useful combination result, based on which we construct a class of OAI Boolean functions in Section 4. In Section 5, these functions are modified to be balanced ones which are also OAI functions, and their algebraic degree, nonlinearity and behavior resisting fast algebraic attacks are studied. Concluding remarks are given in Section 6.

2 Preliminary

Let \mathbb{F}_2 be the binary finite field and \mathbb{F}_2^n be the n -dimensional vector space over \mathbb{F}_2 . An n -variable Boolean function is a mapping from \mathbb{F}_2^n to \mathbb{F}_2 . Denote by \mathbb{B}_n the set of all n -variable Boolean functions. The support of a Boolean function f is defined as

$$\text{supp}(f) = \{\mathbf{x} \in \mathbb{F}_2^n \mid f(\mathbf{x}) = 1\},$$

and the cardinality of it, $\text{wt}(f)$, is called the Hamming weight of f . Furthermore, for another Boolean function $g \in \mathbb{B}_n$, the distance between f and g is defined as $d(f, g) = \text{wt}(f + g)$. When $\text{wt}(f) = 2^{n-1}$, we call f a balanced function. Abusing notations, we also denote the Hamming weight of a vector $\mathbf{v} \in \mathbb{F}_2^n$, i.e. the number of nonzero positions of \mathbf{v} , to be $\text{wt}(\mathbf{v})$. Besides, for an integer u , we denote by $\text{wt}_n(u)$ the number of 1’s in the binary expansion of the reduction of u modulo $(2^n - 1)$ in the complete residue system $\{0, 1, \dots, 2^n - 2\}$. Obviously, $\text{wt}_n(-u) = n - \text{wt}_n(u)$ when $2^n - 1 \nmid u$.

There are several ways to describe a Boolean function such as by its truth table, algebraic normal form (ANF), univariate representation and so on. Each $f \in \mathbb{B}_n$ has a unique ANF of the form

$$f(x_1, \dots, x_n) = \sum_{I \subseteq \{1, 2, \dots, n\}} a_I \prod_{i \in I} x_i, \quad a_I \in \mathbb{F}_2.$$

The algebraic degree of f , $\deg(f)$, is defined to be $\max\{|I| \mid a_I \neq 0\}$. It should be noted that for n -variable balanced Boolean functions, the maximal possible algebraic degree is $(n - 1)$. Boolean functions of degree at most 1 are called affine functions, and the set of all of them are denoted to be \mathbb{A}_n . In order to resist the fast correlation attacks, Boolean functions used in cryptographic systems should have high nonlinearity, where the nonlinearity of a Boolean function f , \mathcal{N}_f , is defined as the minimum distance between f and all affine functions, i.e.

$$\mathcal{N}_f = \min_{a \in \mathbb{A}_n} d(f, a).$$

Walsh transform is a powerful tool in studying Boolean functions. For any $\lambda \in \mathbb{F}_2^n$, the Walsh transform of $f \in \mathbb{B}_n$ at λ is defined by

$$W_f(\lambda) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) + \lambda \cdot x},$$

where “ \cdot ” represents the Euclidean inner product of vectors. Many criteria of f can be described by its Walsh transform such as balancedness, nonlinearity and correlation immunity [2]. For example, we have $W_f(0) = 0$ when f is balanced, and we can equivalently express nonlinearity of f by

$$\mathcal{N}_f = 2^{n-1} - \max_{\lambda \in \mathbb{F}_2^n} |W_f(\lambda)|.$$

As is well known that the finite field \mathbb{F}_{2^n} is isomorphic to \mathbb{F}_2^n through the choice of a basis of \mathbb{F}_{2^n} over \mathbb{F}_2 , hence naturally, the Boolean function f can be represented by a univariate polynomial over \mathbb{F}_{2^n} of the form

$$f(x) = \sum_{i=0}^{2^n-1} f_i x^i.$$

It can be proved that as a Boolean function, the coefficients of f satisfy $f_{2i} = f_i^2$ (subscripts reduced modulo $(2^n - 1)$) for $1 \leq i \leq 2^n - 2$ and $f_0, f_{2^n-1} \in \mathbb{F}_2$. Besides, it is not difficult to deduce that

$$\deg(f) = \max\{\text{wt}_n(i) \mid f_i \neq 0, 0 \leq i \leq 2^n - 1\}.$$

Under univariate representation, the Walsh transform of f at $\lambda \in \mathbb{F}_{2^n}$ can be described as

$$W_f(\lambda) = \sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) + \text{tr}_1^n(\lambda x)},$$

where $\text{tr}_1^n(\cdot)$ is the trace function from \mathbb{F}_{2^n} to \mathbb{F}_2 , i.e. $\text{tr}_1^n(x) = \sum_{i=0}^{n-1} x^{2^i}$ for any $x \in \mathbb{F}_{2^n}$.

When n is even, we can give another formulation of the univariate representation of Boolean functions based on polar decomposition of $\mathbb{F}_{2^n}^*$. Let $n = 2m$. Then $\mathbb{F}_{2^m}^*$ is a cyclic subgroup of $\mathbb{F}_{2^n}^*$. Since $(2^m - 1, \frac{2^n - 1}{2^m - 1}) = (2^m - 1, 2^m + 1) = 1$, there exists a cyclic subgroup U of $\mathbb{F}_{2^n}^*$ of order $2^m + 1$ such that

$$\mathbb{F}_{2^n}^* = \mathbb{F}_{2^m}^* \times U.$$

This is just the polar decomposition of $\mathbb{F}_{2^n}^*$. If we assume α to be a primitive element of \mathbb{F}_{2^n} , then it is obvious that $U = \langle \xi \rangle$ where $\xi = \alpha^{2^m - 1}$. From the polar decomposition we know that any $x \in \mathbb{F}_{2^n}^*$ can be represented as $x = yz$ for some $y \in \mathbb{F}_{2^m}^*$ and $z \in U$. Then we can represent the Boolean function f by

$$f(x) = \begin{cases} f_0 & \text{if } x = 0; \\ f'(x) = f'(y, z) & \text{if } 0 \neq x = yz, y \in \mathbb{F}_{2^m}^*, z \in U, \end{cases}$$

where $f'(x) = \sum_{i=0}^{2^n-2} f'_i x^i$ is the polynomial representation of the map $\mathbb{F}_{2^n}^* \rightarrow \mathbb{F}_2$, $c \mapsto f(c)$ (by Lagrange interpolation). Note that

$$f'(y, z) = \sum_{i=0}^{2^n-2} f'_i (yz)^i = \sum_{j=0}^{2^m-2} \sum_{k=0}^{2^m} f'_{j,k} y^j z^k$$

where for any $0 \leq i \leq 2^n - 2$, $f'_i = f'_{j,k}$ if and only if $\begin{cases} i \equiv j \pmod{2^m - 1} \\ i \equiv k \pmod{2^m + 1} \end{cases}$, i.e. $i \equiv 2^{m-1}((2^m + 1)j + (2^m - 1)k) \pmod{2^n - 1}$ (by the Chinese remainder theorem). Besides,

$$\begin{aligned} f(x) &= f_0(x^{2^n-1} + 1) + f'(x)x^{2^n-1} \\ &= f_0 + f_0 x^{2^n-1} + x^{2^n-1} \sum_{i=0}^{2^n-2} f'_i x^i \\ &\equiv f_0 + (f_0 + f'_0)x^{2^n-1} + \sum_{i=1}^{2^n-2} f'_i x^i \pmod{x^{2^n} + x}, \end{aligned}$$

hence the algebraic degree of f can be expressed as

$$\deg(f) = \begin{cases} \max\{\text{wt}_n(2^{m-1}((2^m + 1)j + (2^m - 1)k)) \mid f'_{j,k} \neq 0\} & \text{if } f_0 + f'_0 = 0; \\ n & \text{if } f_0 + f'_0 \neq 0. \end{cases}$$

That is to say, if the algebraic degree of f is smaller than n , we have $f_0 = f'_0 = f'_{0,0}$ and

$$\begin{aligned}\deg(f) &= \max\{\text{wt}_n(2^{m-1}((2^m + 1)j + (2^m - 1)k)) \mid f'_{j,k} \neq 0\} \\ &= \max\{\text{wt}_n((2^m + 1)j + (2^m - 1)k) \mid f'_{j,k} \neq 0\}.\end{aligned}$$

To finish this section, we recall the definition of algebraic immunity of Boolean functions.

Definition 2.1. Let $f, g \in \mathbb{B}_n$. g is called an annihilator of f if $fg = 0$. The algebraic immunity of f , $\text{AI}(f)$, is defined to be the smallest possible degree of the nonzero annihilators of f or $f + 1$, i.e.

$$\text{AI}(f) = \min_{0 \neq g \in \mathbb{B}_n} \{\deg(g) \mid fg = 0 \text{ or } (f + 1)g = 0\}.$$

3 A combination fact

In this section, we prove a useful combination result about the weight distribution of integers, which will be of key importance in proving the optimal algebraic immunity of the Boolean functions constructed in the following sections.

Lemma 3.1. Let $n = 2m$. Then for any $0 \leq j \leq 2^m - 2$, $1 \leq k \leq 2^m$, we have

$$\text{wt}_n((2^m + 1)(2^m - 1 - j) + (2^m - 1)k) = n - \text{wt}_n((2^m + 1)j + (2^m - 1)k).$$

Proof. Obviously,

$$\begin{aligned}2^m[2^m(j - k) + (j + k)] &\equiv 2^n(j - k) + 2^m(j + k) \\ &\equiv 2^m(j + k) + (j - k) \pmod{2^n - 1},\end{aligned}$$

and thus

$$\begin{aligned}\text{wt}_n(2^m(j - k) + (j + k)) &= \text{wt}_n(2^m[2^m(j - k) + (j + k)]) \\ &= \text{wt}_n(2^m(j + k) + (j - k)) \\ &= \text{wt}_n((2^m + 1)j + (2^m - 1)k).\end{aligned}$$

Then we get

$$\begin{aligned}
wt_n((2^m + 1)(2^m - 1 - j) + (2^m - 1)k) &= wt_n(2^n - 1 - (2^m + 1)j + (2^m - 1)k) \\
&= n - wt_n((2^m + 1)j - (2^m - 1)k) \\
&= n - wt_n(2^m(j - k) + (j + k)) \\
&= n - wt_n((2^m + 1)j + (2^m - 1)k).
\end{aligned}$$

□

Proposition 3.2. *Let $n = 2m$. For any $0 \leq k \leq 2^m$, define*

$$S_k = \{j \in \mathbb{Z}/(2^m - 1)\mathbb{Z} \mid wt_n((2^m + 1)j + (2^m - 1)k) < m\}.$$

Then $|S_k| \leq 2^{m-1}$, $0 \leq k \leq 2^m$. Moreover, “=” holds if and only if m is odd and $k = 0$.

Proof. Consider the case $k = 0$ firstly. Since

$$\begin{aligned}
S_0 &= \{j \in \mathbb{Z}/(2^m - 1)\mathbb{Z} \mid wt_n((2^m + 1)j) < m\} \\
&= \{j \in \mathbb{Z}/(2^m - 1)\mathbb{Z} \mid wt_n(2^m j + j) < m\} \\
&= \{j \in \mathbb{Z}/(2^m - 1)\mathbb{Z} \mid wt_m(j) < \frac{m}{2}\},
\end{aligned}$$

it is easy to get

$$\begin{aligned}
|S_0| &= \begin{cases} \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \binom{m}{i} & \text{if } m \text{ is odd;} \\ \sum_{i=0}^{\frac{m}{2}-1} \binom{m}{i} & \text{if } m \text{ is even} \end{cases} \\
&= \begin{cases} 2^{m-1} & \text{if } m \text{ is odd;} \\ 2^{m-1} - \frac{1}{2} \binom{m}{m/2} & \text{if } m \text{ is even,} \end{cases}
\end{aligned}$$

which implies that $|S_0| < 2^{m-1}$ when m is even.

Now we consider the case $1 \leq k \leq 2^m$. Define the set

$$T_k = \{j \in \mathbb{Z}/(2^m - 1)\mathbb{Z} \mid wt_n((2^m + 1)j + (2^m - 1)k) > m\}.$$

From Lemma 3.1 we have for any $0 \leq j \leq 2^m - 2$,

$$wt_n((2^m + 1)(2^m - 1 - j) + (2^m - 1)k) = n - wt_n((2^m + 1)j + (2^m - 1)k),$$

thus $|S_k| = |T_k|$. On the other hand, since

$$\begin{aligned} \text{wt}_n((2^m - 1)k) &= \text{wt}_n(2^m(2^m k - k)) \\ &= \text{wt}_n(k - 2^m k) \\ &= n - \text{wt}_n((2^m - 1)k), \end{aligned}$$

i.e. $\text{wt}_n((2^m - 1)k) = \frac{n}{2} = m$, we know that $0 \notin S_k$ and $0 \notin T_k$, which implies

$$|S_k| + |T_k| \leq 2^m - 2$$

as $S_k \cap T_k = \emptyset$. Then it follows that $|S_k| \leq 2^{m-1} - 1 < 2^{m-1}$. \square

4 A class of unbalanced OAI Boolean functions

In this section, based on polar decomposition of $\mathbb{F}_{2^m}^*$ and the combination results in Section 3, we construct a new class of OAI Boolean functions.

Construction 4.1. *Let $n = 2m$. Let β be a primitive element of \mathbb{F}_{2^m} and U be the cyclic group defined in Section 2. Set $\Delta = \{1, \beta, \beta^2, \dots, \beta^{2^{m-1}-1}\}$. Define an n -variable Boolean function f by setting*

$$\text{supp}(f) = \Delta \times U.$$

Theorem 4.2. *Let f be the Boolean function defined in Construction 4.1. Then f has optimal algebraic immunity.*

Proof. From the definition of algebraic immunity, it suffices to prove that there is no nonzero annihilator with degree smaller than m of both f and $f + 1$.

Suppose $g \neq 0$ is an annihilator of f with algebraic degree smaller than m . Assume

$$g(x) = \begin{cases} g'(y, z) & \text{if } 0 \neq x = yz, y \in \mathbb{F}_{2^m}^*, z \in U; \\ g_0 & \text{if } x = 0, \end{cases}$$

where $g'(y, z) = \sum_{j=0}^{2^m-2} \sum_{k=0}^{2^m} g_{j,k} y^j z^k$, $g_{j,k} \in \mathbb{F}_{2^n}$, $g_0 \in \mathbb{F}_2$. Then

$$g'(y, z) = \sum_{k=0}^{2^m} \left(\sum_{j=0}^{2^m-2} g_{j,k} y^j \right) z^k = \sum_{k=0}^{2^m} g_k(y) z^k = 0$$

for all $z \in U$ and $y \in \Delta$, where $g_k(y) = \sum_{j=0}^{2^m-2} g_{j,k} y^j$. For any fixed $y_0 \in \Delta$, since $g'(y_0, z)$ has $(2^m + 1)$ zeros, we conclude that $g_k(y) = 0$ for any $y \in \Delta$, $0 \leq k \leq 2^m$.

On the one hand, from the definition of BCH code [17], we know that for $0 \leq k \leq 2^m$, $(g_{0,k}, g_{1,k}, g_{2,k}, \dots, g_{2^m-2,k})$ is a codeword of some BCH code over \mathbb{F}_{2^n} of length $(2^m - 1)$ with elements in Δ as zeros. Thus based on the BCH bound, the Hamming weight of a nonzero codeword should be greater than or equal to $(2^{m-1} + 1)$, i.e.

$$\text{wt}(g_{0,k}, g_{1,k}, g_{2,k}, \dots, g_{2^m-2,k}) \geq 2^{m-1} + 1.$$

On the other hand, since $\deg(g) < m$, we have $g_{j,k} = 0$ if $\text{wt}_n((2^m + 1)j + (2^m - 1)k) \geq m$. Form Proposition 3.2, we know that $|S_k| \leq 2^{m-1}$. That is

$$\text{wt}(g_{0,k}, g_{1,k}, g_{2,k}, \dots, g_{2^m-2,k}) \leq 2^{m-1},$$

which leads to a contradiction. Hence $g = 0$.

Next, we consider the function $f + 1$. Note that

$$\text{supp}(f + 1) = \Delta' \times U \cup \{0\}$$

where $\Delta' = \{\beta^{2^{m-1}}, \beta^{2^{m-1}+1}, \dots, \beta^{2^m-2}\}$. Similar to the proof with respect to f , we let g be now a nonzero annihilator of $f + 1$ with algebraic degree smaller than m . We can deduce from the BCH bound that, for any $0 \leq k \leq 2^m$, the vector $(g_{0,k}, g_{1,k}, g_{2,k}, \dots, g_{2^m-2,k})$ has weight at least 2^{m-1} since $|\Delta'| = 2^{m-1} - 1$. By Proposition 3.2, we know that when m is even, the weight of $(g_{0,k}, g_{1,k}, \dots, g_{2^m-2,k})$ is smaller than 2^{m-1} , thus a contradiction follows and $g = 0$. When m is odd, we have $(g_{0,k}, g_{1,k}, g_{2,k}, \dots, g_{2^m-2,k}) = (0, 0, \dots, 0)$ for $1 \leq k \leq 2^m$. Since $|S_0| = 2^{m-1}$, we get $\text{wt}((g_{0,0}, g_{1,0}, \dots, g_{2^m-2,0})) = 2^{m-1}$, which implies that $g_{0,0} = g_0 = 1$. However, this contradicts the fact that $0 \in \text{supp}(f + 1)$. We also have $g = 0$.

To summarize, we know that f has optimal algebraic immunity. \square

Remark 4.3. From the proof of Theorem 4.2, it is easy to see that if we replace the set Δ in Construction 4.1 by $\{\beta^s, \beta^{s+1}, \dots, \beta^{s+2^{m-1}-1}\}$ for any $0 \leq s \leq 2^m - 2$, we can also obtain Boolean functions with optimal algebraic immunity.

It is direct to find that the weight of the function in Construction 4.1 is $(2^{n-1} + 2^{m-1})$, which is bigger than that of balanced functions. Thus we do not talk about their further properties since they are not of applicable interest.

5 Balanced functions with optimal algebraic immunity and other good properties

In this section, we modify the functions in Construction 4.1 to be balanced ones which maintain optimal algebraic immunity by changing some points between their supports and zeros. Furthermore, we study in detail properties of these balanced functions such as their algebraic degree, nonlinearity and immunity against fast algebraic attacks.

Construction 5.1. *Let $n = 2m$. Let α be a primitive element of \mathbb{F}_{2^n} and $\beta = \alpha^{2^m+1}$, $\xi = \alpha^{2^m-1}$. Set $\Gamma = \{\beta, \beta^2, \dots, \beta^{2^{m-1}-1}\}$. Define an n -variable Boolean function F by setting*

$$\text{supp}(F) = (\Gamma \times U) \cup (\{1\} \times \{1, \xi, \dots, \xi^{2^m-1}\}).$$

Theorem 5.2. *Let F be the Boolean function defined in Construction 5.1. Then F is balanced and has optimal algebraic immunity.*

Proof. It is obvious that $\text{wt}(F) = (2^{m-1} - 1) \times (2^m + 1) + 2^{m-1} + 1 = 2^{n-1}$, so F is balanced.

The proof of optimal algebraic immunity of F is similar to that of Theorem 4.2. Suppose g is a nonzero annihilator of F with algebraic degree smaller than m , and assume

$$g(x) = \begin{cases} g'(y, z) & \text{if } 0 \neq x = yz, y \in \mathbb{F}_{2^m}^*, z \in U; \\ g_0 & \text{if } x = 0, \end{cases}$$

where $g'(y, z) = \sum_{j=0}^{2^m-2} \sum_{k=0}^{2^m} g_{j,k} y^j z^k$, $g_{j,k} \in \mathbb{F}_{2^n}$, $g_0 \in \mathbb{F}_2$. Since $\{\beta, \beta^2, \dots, \beta^{2^{m-1}-1}\} \times$

$U \subseteq \text{supp}(f)$, by the BCH bound and Proposition 3.2, we get for $k > 0$, $(g_{0,k}, g_{1,k}, \dots, g_{2^m-2,k}) = (0, 0, \dots, 0)$. Then $g'(y, z)$ turns to $g'(y, z) =$

$\sum_{j=0}^{2^m-2} g_{j,0} y^j$. Besides, as $\{1\} \times \{1, \xi, \dots, \xi^{2^m-1}\} \subseteq \text{supp}(f)$, we have

$$g'(1, z) = \sum_{j=0}^{2^m-2} g_{j,0} 1^j = 0,$$

which means that $\{1, \beta, \beta^2, \dots, \beta^{2^m-1-1}\}$ are zeros of certain BCH code containing $(g_{0,0}, g_{1,0}, \dots, g_{2^m-2,0})$ as a codeword. Using the BCH bound and Proposition 3.2 again, we obtain a contradiction. Thus F has no nonzero annihilator with degree smaller than m . With respect to $F + 1$, the proof procedure is almost the same.

Finally, we conclude that the Boolean function F has optimal algebraic immunity. \square

Remark 5.3. *From the proof of Theorem 5.2, it is not difficult to see that we can also set $\text{supp}(F) = (\{1, \beta, \dots, \beta^{2^m-1-2}\} \times U) \cup (\{\beta^{2^m-1-1}\} \times \{1, \xi, \dots, \xi^{2^m-1}\})$ to obtain balanced Boolean functions with optimal algebraic immunity.*

5.1 Polynomial representation and algebraic degree

In the following, we compute the univariate representation of the OAI Boolean function F in Construction 5.1 and deduce its algebraic degree.

By the Chinese remainder theorem, we can write the support of F in the form

$$\begin{aligned} \text{supp}(F) = & \{ \alpha^{2^{m-1}((2^m+1)j+(2^m-1)k)} \mid 1 \leq j \leq 2^{m-1} - 1, 0 \leq k \leq 2^m \} \\ & \cup \{ \alpha^{2^{m-1}(2^m-1)k} \mid 0 \leq k \leq 2^{m-1} \}. \end{aligned}$$

For simplicity, we distinguish the integer $2^{m-1}((2^m+1)j+(2^m-1)k)$ reduced modulo (2^n-1) with a pair (j, k) where $0 \leq j \leq 2^m-2, 0 \leq k \leq 2^m$. It is easy to find that

$$\begin{aligned} (j+1, k+1) &= 2^{m-1}((2^m+1)(j+1) + (2^m-1)(k+1)) \\ &= 2^{m-1}((2^m+1)j + (2^m-1)k) + 1 \\ &= (j, k) + 1, \end{aligned}$$

and

$$(j, k-2) = 2^{m-1}((2^m+1)j + (2^m-1)(k-2))$$

$$\begin{aligned}
&= 2^{m-1}((2^m + 1)j + (2^m - 1)k) + (2^m - 1) \\
&= (j, k) + (2^m - 1).
\end{aligned}$$

Using these properties, we can derive that the support of F is just

$$\begin{aligned}
\text{supp}(F) &= \{\alpha^{(j,k)} \mid 1 \leq j \leq 2^{m-1} - 1, 0 \leq k \leq 2^m\} \cup \{\alpha^{(0,k)} \mid 0 \leq k \leq 2^{m-1}\} \\
&= \{\alpha^{l(2^m-1)+r} \mid 0 \leq l \leq 2^m, 1 \leq r \leq 2^{m-1} - 1\} \\
&\quad \cup \{\alpha^{2^{m-1}(2^m-1)k} \mid 0 \leq k \leq 2^{m-1}\}.
\end{aligned}$$

Then the coefficients of the function f' whose support is the first part of $\text{supp}(F)$ can be described explicitly, i.e. for $0 < i < 2^n - 1$,

$$\begin{aligned}
f'_i &= \sum_{l=0}^{2^m} \sum_{j=l(2^m-1)+1}^{l(2^m-1)+2^{m-1}-1} (\alpha^{-i})^j \\
&= \sum_{l=0}^{2^m} \frac{(\alpha^{-i})^{1+l(2^m-1)}(1 - (\alpha^{-i})^{2^{m-1}-1})}{1 - \alpha^{-i}} \\
&= \frac{\alpha^{-i}(1 - \alpha^{-i(2^{m-1}-1)})}{1 - \alpha^{-i}} \sum_{l=0}^{2^m} \alpha^{-il(2^m-1)} \\
&= \begin{cases} 0 & \text{if } 2^m + 1 \nmid i; \\ \frac{\alpha^{-i}(1 - \alpha^{-i(2^{m-1}-1)})}{1 - \alpha^{-i}} & \text{if } 2^m + 1 \mid i. \end{cases}
\end{aligned}$$

Similarly, the coefficients of f'' whose support is the second part of $\text{supp}(F)$ are that, for $0 < i < 2^n - 1$,

$$\begin{aligned}
f''_i &= \sum_{k=0}^{2^{m-1}} (\alpha^{-i})^{2^{m-1}(2^m-1)k} \\
&= \begin{cases} \frac{1 - \alpha^{-i2^{m-1}(2^m-1)(2^{m-1}+1)}}{1 - \alpha^{-i2^{m-1}(2^m-1)}} & \text{if } 2^m + 1 \nmid i; \\ 1 & \text{if } 2^m + 1 \mid i. \end{cases}
\end{aligned}$$

It is obvious that, if we assume $F(x) = \sum_{i=0}^{2^n-1} F_i x^i$, then $F_i = f'_i + f''_i$ for $1 \leq i \leq 2^n - 2$, $F_0 = 0$ (since $F(0) = 0$) and $F_{2^n-1} = 0$ (since F is balanced). Hence we can give the univariate representation of F .

Theorem 5.4. *Let F be the n -variable Boolean function defined in Construction 5.1. Then the univariate representation of F is*

$$F(x) = \sum_{i=1}^{2^n-2} F_i x^i,$$

where

$$F_i = \begin{cases} \frac{1 - \alpha^{-i2^{m-1}(2^m-1)(2^{m-1}+1)}}{1 - \alpha^{-i2^{m-1}(2^m-1)}} & \text{if } 2^m + 1 \nmid i; \\ 1 + \frac{\alpha^{-i}(1 - \alpha^{-i(2^{m-1}-1)})}{1 - \alpha^{-i}} & \text{if } 2^m + 1 \mid i. \end{cases}$$

Hence the algebraic degree of F is $(n - 1)$, which is optimal for balanced Boolean functions.

5.2 Nonlinearity

To determine the lower bound of the nonlinearity of the Boolean functions in Construction 5.1, we need some necessary backgrounds.

Definition 5.5 ([11]). *Let $a \in \mathbb{F}_{2^m}$. The binary complete Kloosterman sum is defined as*

$$\mathcal{K}(a) = \sum_{x \in \mathbb{F}_{2^m}} (-1)^{\text{tr}_1^m(1/x + ax)}.$$

Lemma 5.6 ([10]). *Let $a \in \mathbb{F}_{2^m}^*$ and U be the cyclic group defined in Section 2. Then*

$$\sum_{z \in U} (-1)^{\text{tr}_1^n(az)} = 1 - \mathcal{K}(a).$$

Lemma 5.7 ([20]). *Let β be a primitive element of \mathbb{F}_{2^m} . Let*

$$\Delta_s = \{\beta^s, \beta^{s+1}, \dots, \beta^{2^{m-1}+s-1}\}$$

where $0 \leq s < 2^m - 1$ is an integer. Then

$$\left| \sum_{\gamma \in \Delta_s} (\mathcal{K}(\gamma) - 1) \right| < \left(\frac{\ln 2}{\pi} + 0.42 \right) 2^m + 1.$$

Theorem 5.8. *Let F be the Boolean function defined in Construction 5.1. Then*

$$\mathcal{N}_F > 2^{n-1} - \left(\frac{\ln 2}{\pi}m + 0.92\right)2^m - 1.$$

Proof. We denote the set $\{1, \xi, \dots, \xi^{2^{m-1}}\}$ by Λ . Obviously, $W_F(0) = 0$ since F is balanced.

For any $a \in \mathbb{F}_{2^n}^*$, we assume $a = a_1 a_2$ where $a_1 \in \mathbb{F}_{2^m}^*$, $a_2 \in U$. By Lemma 5.6, we have

$$\begin{aligned} W_F(a) &= -2 \sum_{x \in \text{supp}(F)} (-1)^{\text{tr}_1^n(ax)} \\ &= -2 \left[\sum_{y \in \Gamma} \sum_{z \in U} (-1)^{\text{tr}_1^n(a_1 y a_2 z)} + \sum_{z \in \Lambda} (-1)^{\text{tr}_1^n(a_1 a_2 z)} \right] \\ &= -2 \left[\sum_{y \in \Gamma'} (1 - \mathcal{K}(a_1 y)) - \sum_{z \in U} (-1)^{\text{tr}_1^n(a_1 z)} + \sum_{z \in \Lambda'} (-1)^{\text{tr}_1^n(a_1 z)} \right] \\ &= -2 \left[\sum_{y \in \Gamma'} (1 - \mathcal{K}(a_1 y)) - \sum_{z \in U \setminus \Lambda'} (-1)^{\text{tr}_1^n(a_1 z)} \right], \end{aligned}$$

where $\Lambda' = \{a_2, a_2 \xi, \dots, a_2 \xi^{2^{m-1}}\}$, $\Gamma' = \{1, \beta, \dots, \beta^{2^{m-1}-1}\}$.

Since $a_1 \in \mathbb{F}_{2^m}^*$, it can be represented as $a_1 = \beta^s$ for some $0 \leq s \leq 2^m - 2$. Then

$$\sum_{y \in \Gamma'} (\mathcal{K}(a_1 y) - 1) = \sum_{\gamma \in \Delta_s} (\mathcal{K}(\gamma) - 1),$$

where $\Delta_s = \{\beta^s, \beta^{s+1}, \dots, \beta^{s+2^{m-1}-1}\}$. By Lemma 5.7, we know that

$$\left| \sum_{y \in \Gamma'} (1 - \mathcal{K}(a_1 y)) \right| < \left(\frac{\ln 2}{\pi}m + 0.42 \right) 2^m + 1.$$

Therefore,

$$|W_F(a)| < 2 \left[\left(\frac{\ln 2}{\pi}m + 0.42 \right) 2^m + 1 + 2^{m-1} \right].$$

Finally we get

$$\mathcal{N}_F = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_{2^n}} |W_f(a)|$$

$$\begin{aligned}
&> 2^{n-1} - \left(\frac{\ln 2}{\pi} m + 0.42 \right) 2^m - 2^{m-1} - 1 \\
&= 2^{n-1} - \left(\frac{\ln 2}{\pi} m + 0.92 \right) 2^m - 1.
\end{aligned}$$

□

In fact, the lower bound in Theorem 5.8 is not satisfactory at all since we have used the naive estimation

$$\left| \sum_{z \in U \setminus \Lambda'} (-1)^{\text{tr}_1^n(a_1 z)} \right| \leq 2^{m-1}$$

in the proof. Hence it is not so safe to say that the function F has good nonlinearity. Nevertheless, for these n we can compute the exact value of nonlinearity, it appears good.

Denote by \mathcal{N}_{C-F} , \mathcal{N}_{T-C-T} and \mathcal{N}_F the exact values of nonlinearity of the Carlet-Feng functions [5], the Tang-Carlet-Tang functions [20] and the functions in Construction 5.1 respectively. By a Magma program, we investigate the exact values of nonlinearity for small number of variables under the choice of the default primitive element of \mathbb{F}_{2^n} in Magma system. The results are displayed in Table 1. By the comparison, we find our functions almost play as well as the Carlet-Feng and Tang-Carlet-Tang functions.

Table 1: Comparison of the exact values of Nonlinearity with some known constructions

n	\mathcal{N}_{C-F}	\mathcal{N}_{T-C-T}	\mathcal{N}_F	$2^{n-1} - 2^{\frac{n}{2}-1}$
4	4	4	4	6
6	24	22	22	28
8	112	108	108	120
10	478	476	474	496
12	1970	1982	1976	2016
14	8036	8028	8026	8128
16	32530	32508	32498	32540
18	130442	130504	130484	130812
20	523154	523144	523122	523776

To obtain better estimation of the nonlinearity of the functions in Construction 5.1, the key difficulty is to estimate such exponential sums as

$$\Phi_s = \sum_{x \in \{\xi^s, \xi^{s+1}, \dots, \xi^{s+2^m-1}-1\}} (-1)^{\text{tr}_1^n(cx)}$$

for any $0 \leq s \leq 2^m$ and $c \in \mathbb{F}_{2^m}^*$, where ξ is a generator of the cyclic group U . Unfortunately, the standard technique of using Gauss sums would not work for this kind of incomplete exponential sums over finite fields. Maybe more advanced number theoretic tools should be introduced to overcome this difficulty. Though we have not found them up to present, we conjecture that $|\Phi_s| = O(2^{\frac{m}{2}})$.

5.3 Immunity against fast algebraic attacks

The property of optimal algebraic immunity is a necessary but not sufficient condition for a Boolean function because of the existence of fast algebraic attacks. In this subsection, we analyze the ability of the Boolean functions in Construction 5.1 against fast algebraic attacks.

An n -variable Boolean function f is optimal with respect to fast algebraic attacks if for any pair of integers (e, d) such that $e + d < n$ and $e < n/2$, there do not exist a function $g \neq 0$ of algebraic degree at most e such that fg has degree at most d [6]. Armknecht et.al. proposed an efficient algorithm [1] to determine the existence of g and h with corresponding degrees. Based on Algorithm 2 in [1], we investigate the behavior of the functions in Construction 5.1 against fast algebraic attacks for small number of variables by a Magma program.

We choose the default primitive element of \mathbb{F}_{2^n} in the Magma system. For even n ranging from 4 to 14 and $e < \frac{n}{2}$, we can find the pairs (e, d) with $e + d \geq n - 1$, but the pairs (e, d) such that $e + d \leq n - 2$ have never been observed. That implies that the functions in Construction 5.1 have good immunity to fast algebraic attacks though they are not optimal.

6 Concluding remarks

In this paper, based on polar decomposition of multiplicative groups of quadratic extensions of finite fields, we construct two classes of algebraic immunity optimal Boolean functions. We find that the second class of Boolean

functions possess almost all the necessary properties to be used as filter functions in stream ciphers.

In fact, in the proof of Theorem 5.2, no property of the set $\Lambda = \{1, \xi, \dots, \xi^{2^m-1}\}$ has been used except the cardinality of it. Therefore, we can construct balanced OAI Boolean functions by setting $\text{supp}(F) = (\Gamma \times U) \cup (\{1\} \times \Lambda')$ for any subset Λ' of U satisfying $|\Lambda'| = 2^{m-1} + 1$. Then we have more opportunities to get balanced OAI Boolean functions with high nonlinearity. However, univariate representations and algebraic degrees of functions constructed using Λ' with no special properties would be difficult to describe.

References

- [1] Armknecht F., Carlet C., Gabort P., et al.: Efficient Computation of algebraic immunity for algebraic and fast algebraic attack. In: Vaudenay, S. (eds.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 147–164. Springer, Heidelberg (2006).
- [2] Carlet C.: Boolean functions for cryptography and error correcting codes. In: Crama Y., Hammer P. (eds.) Monography Boolean Methods and Models. Cambridge University Press, London (2010).
- [3] Carlet C.: On a weakness of the Tu-Deng function and its repair. Cryptology ePrint Archive, Report 2009/606 (2009), <http://eprint.iacr.org/>
- [4] Carlet C., Dalai D., Gupta K., Maitra S.: Algebraic Immunity for Cryptographically Significant Boolean Functions, Analysis and Construction. IEEE Trans. Inform. Theory 52, 3105–3121 (2006).
- [5] Carlet C., Feng K.: An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity. In: Pieprzyk J. (eds.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 425–440. Springer, Heidelberg (2008).
- [6] Courtois N.: Fast algebraic attacks on stream ciphers with linear feedback. In: Boneh D. (eds.) CRYPTO 2003. LNCS, vol. 2729, pp. 176–194. Springer, Heidelberg (2003).
- [7] Courtois N., Meier W.: Algebraic attack on stream ciphers with linear feedback. In: Biham E. (eds.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 345–359. Springer, Heidelberg (2003).

- [8] Dalai D., Maitra S., Sarkar S.: Basic theory in construction of boolean functions with maximum possible annihilator immunity. *Des. Codes Cryptogr.* 40, 41–58 (2006).
- [9] Jin Q., Liu Z., Wu B., Zhang X.: A general conjecture similar to T-D conjecture and its application in constructing Boolean functions with optimal algebraic immunity. *Cryptology eprint Archive*, Report 2011/515 (2011), <http://eprint.iacr.org/>
- [10] Leander G.: Monomial bent functions. *IEEE Trans. Inform. Theory* 52, pp. 738–743 (2006).
- [11] Lidl H., Niederreiter R.: *Finite Field*. Cambridge University Press, London (1997).
- [12] Li N., Qi W.: Construction and analysis of boolean functions of $2t + 1$ variables with maximum algebraic immunity. In: Lai X., Chen K. (eds.) *ASIACRYPT 2006*. LNCS, vol. 4284, pp. 84–98. Springer, Heidelberg (2006).
- [13] Li N., Qu L., Qi W., et al.: On the construction of Boolean functions with optimal algebraic immunity. *IEEE Trans. Inform. Theory* 54, 1330–1334 (2008).
- [14] Liu M., Zhang Y., Lin D.: Perfect algebraic immune functions. In: Wang X., Sako K. (eds.) *ASIACRYPT 2012*. LNCS, vol. 7658, pp. 172–189. Springer, Heidelberg (2012).
- [15] Liu M., Zhang Y., Lin D.: On the immunity of Boolean functions against fast algebraic attacks using bivariate polynomial representation. *Cryptology eprint Archive*, Report 2012/498 (2012), <http://eprint.iacr.org/>
- [16] Lou Y., Han H., Tang C., Xu M.: Constructing Vectorial Boolean Functions With High Algebraic Immunity Based on Group Decomposition. *Cryptology eprint Archive*, Report 2012/335 (2012), <http://eprint.iacr.org/>
- [17] MacWilliams F., Sloane N.: *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam (1977).

- [18] Meier W., Pasalic E., Carlet C.: Algebraic attacks and decomposition of boolean functions. In: Cachin, C., Camenisch J. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 474–491. Springer, Heidelberg (2004).
- [19] Mesnager S.: A new class of bent and hyper-bent Boolean functions in polynomial forms. *Des. Codes Cryptogr.* 59, 265–279 (2011).
- [20] Tang D., Carlet C., Tang X.: Highly nonlinear Boolean functions with optimum algebraic immunity and good behavior against fast algebraic attacks. *IEEE Trans. Inform. Theory* 59, 653–664 (2013).
- [21] Tu Z., Deng Y.: A conjecture about binary strings and its applications on constructing Boolean functions with optimal algebraic immunity. *Des. Codes Cryptogr.* 60(1), 1–14 (2011).